

Prototipo de recomendación inteligente de anuncios en portal de ventas online: una experiencia de vinculación entre institutos de investigación y empresa

Sebastián Vallejos¹, Luis Berdun¹, Marcelo Armentano¹, Silvia Schiaffino¹
Sandra González Císaro², Oscar Nigro², Ignacio Cuesta³ y Leonardo Balduzzi³

¹ ISISTAN (CONICET - UNCPBA) - Campus Universitario, Paraje Arroyo Seco, Tandil, Argentina

{sebastian.vallejos, luis.berdun, marcelo.armentano, silvia.schiaffino}@isistan.unicen.edu.ar

² INTIA, Fac. Cs. Exactas - Campus Universitario, Paraje Arroyo Seco, Tandil, Argentina

{sagonci, onigro}@exa.unicen.edu.ar

³ Región Global, Tandil

{ignacio.cuesta, leonardo.balduzzi}@regionglobal.com

Resumen. El objetivo del proyecto consistía originalmente en desarrollar un prototipo de asistente inteligente que pudiera acercar a las partes del comercio electrónico (comprador - vendedor) mediante una caracterización inteligente de cada usuario junto con sus preferencias a corto plazo y características geoespaciales que lo definan. El objetivo consistía en brindar alertas inteligentes sobre artículos que puedan interesarles. El proyecto considera como aspectos fundamentales para su fácil adopción la posibilidad de adquirir el conocimiento del usuario de manera inteligente y no intrusiva. A partir del trabajo en conjunto entre la empresa y los investigadores se fueron adaptando los distintos objetivos a fin de poder abarcar las necesidades efectivas de la empresa.

1 Introducción

El Comercio electrónico ha evolucionado hasta convertirse en la actualidad una de las principales opciones de venta. Según la Cámara Argentina de Comercio Electrónico (CACE) en el año 2018 la facturación del comercio electrónico creció en un 47% respecto al año 2017 logrando una facturación total de \$229.760 millones de pesos. En este contexto, los portales de comercio electrónico como Region20 se han convertido en una fuente fundamental para brindar al usuario una forma de poder comercializar sus productos (nuevos o usados) o servicios. En este tipo de portales, tanto usuarios particulares como comercios locales ven la posibilidad de acceder a esta forma de comercialización de productos o servicios a un costo reducido.

Para lograr una efectiva simbiosis entre el comprador y el vendedor es necesario poder llegar al comprador con los productos que este requiere, en el tiempo que los requiere y con la posibilidad de acceder a estos en el menor tiempo posible. Para

poder realizar esto es necesario conocer el perfil del comprador, es decir tanto sus características generales que definen sus hábitos de compra, como así como también sus necesidades puntuales en el corto plazo. Para la realización de una venta efectiva también debe superarse el umbral de espera, es decir la posibilidad de unir al comprador con el producto en el menor tiempo posible. La logística en la entrega de una venta es uno de los principales desafíos que enfrenta el comercio electrónico.

En este contexto, el proyecto abordado tuvo por objetivo realizar un prototipo que pueda acercar las partes (comprador-vendedor) mediante una caracterización inteligente de cada usuario junto con sus preferencias a corto plazo. De esta forma, se buscó poder brindar alertas inteligentes sobre artículos que puedan interesarles y, de ser posible, en zonas de proximidad geográficas. El proyecto considera como aspectos fundamentales para su fácil adopción la posibilidad de adquirir el conocimiento del usuario de forma no intrusiva y con información pre-existente.

Este trabajo es parte del resultado de un proyecto Fase Cero financiado por la Fundación Sadosky entre dos institutos de investigación de la Facultad de Ciencias Exactas de la UNCPBA, ISISTAN (CONICET - UNCPBA) e INTIA, y la empresa Región Global. Esta empresa posee un sitio de comercio electrónico zonal denominado Región 20¹.

El artículo se encuentra organizado de la siguiente manera. En la sección 2 se mencionan los antecedentes del proyecto. En la sección se detallan los objetivos. En la sección 4 se describen las etapas del proyecto. En la sección 5 se describen los resultados alcanzados. Finalmente en la sección 6 se exponen las conclusiones.

2 Antecedentes

El problema del perfilado y clasificación de clientes ha sido estudiado anteriormente. Distintos enfoques han sido utilizados para el perfilado de usuarios en e-commerce utilizando información sobre la navegación y patrones de conducta de los clientes ya sea para clasificar a los clientes [1,2,3] o para poder predecir su próxima compra [4]. En este contexto es importante identificar a los usuarios dentro del e-commerce para realizar acciones tales como la recomendación de anuncios y poder convertirlos de usuarios a clientes.

La recomendación de información contextualizada a partir de la geolocalización del usuario es un área que ha recibido especial atención en la comunidad de sistemas de recomendación en los últimos años. Este tipo de campañas publicitarias, complementan las campañas por medios tradicionales (televisión, radio, Internet) permitiendo la generación de campañas publicitarias personalizadas de acuerdo a los intereses del usuario, su ubicación actual, sus necesidades del momento y los dispositivos que utilizan.

¹ <https://www.region20.com.ar/>

En [5], por ejemplo, se presenta una infraestructura para la recomendación de publicidades de actividades tanto comerciales como no comerciales. Esta arquitectura busca minimizar las interacciones del usuario intentando aprender lo más posible acerca de sus hábitos de manera implícita. Esta arquitectura, hace uso de las antenas de telefonía celular para deducir la ubicación del usuario, con una precisión mucho inferior a la que podría obtenerse utilizando las antenas GPS. La información utilizada para caracterizar tanto a los usuarios como a los anuncios a recomendar incluyen las siguientes características: categoría (venta por mayor y menor, arte y entretenimiento, etc), día (días laborables, fines de semana), franja horaria (antes de las 17 hs y después de las 17 hs), lugar (exterior, interior y formal, interior e informal), costo (gratis, con costo), ejecutor (celebridades, otros). El modelo predictivo utiliza una red neuronal de dos capas y estereotipos de usuarios para recomendar una lista acotada y ordenada de anuncios a los usuarios.

En el trabajo presentado en [6] también se utiliza un aprendizaje de preferencias de los usuarios a partir de sus actividades online. Este enfoque, utiliza tanto la información de los productos por los que el usuario navega en un sitio de comercio electrónico, como los comentarios dejados por los clientes acerca de los productos adquiridos. Recientemente, en [7] se concluye que en la publicidad basada en la ubicación, el diseño publicitario más efectivo se basa en ofertas ordenadas por distancia. Los resultados muestran que la distancia geográfica tiene un impacto negativo en el comportamiento de elección de cupones de ofertas (es decir, cuanto más alejados están los consumidores de una tienda, menos probabilidades tienen de elegir un cupón relacionado), lo que puede explicarse por los costos de transporte. Sin embargo, también se observa que los consumidores aceptan mayores costos de transporte para recibir descuentos más atractivos.

En cuanto a antecedentes del grupo de investigación en temas relacionados al proyecto, el grupo de investigación posee experiencia en la identificación de información de geolocalización a partir de textos planos. El trabajo desarrollado en [8; 9,10] obtiene automáticamente información de geolocalización de los incidentes de tránsito que afectan a una determinada región, a partir del conocimiento extraído de publicaciones en redes sociales. La solución ad-hoc desarrollada en este proyecto servirá de puntapié inicial para analizar e identificar información acerca de la ubicación de los productos o servicios ofrecidos en diferentes anuncios publicitarios cuando no se cuente explícitamente con esta información.

Relativo al aprendizaje de perfiles de rutinas de transporte de un usuario, en [11] se presentó un enfoque de un asistente inteligente para asistir a los ciudadanos en el traslado urbano. El prototipo desarrollado permite aprender de manera no intrusiva la rutina de un usuario y, a partir de esto, realizar sugerencias que le permitan evitar problemas en su traslado (congestionamientos, manifestaciones, accidentes, calles cortadas, etc). De manera similar, el enfoque presentado en [12] aprende los intereses de los usuarios relativos a atracciones turísticas. Los perfiles de usuario se infiere automáticamente en base a los recorridos realizados por los usuarios y el tiempo de permanencia en cada punto de interés (POI). En base a esta información histórica, cuando un usuario visita una nueva ciudad, se le ofrecen atracciones que pueden serle

de interés. Los resultados de estos trabajos pueden sustentar parte del desarrollo del proyecto presentado ya que, dada la posición actual del usuario, sería posible predecir a dónde se va a dirigir a continuación de acuerdo a su rutina diaria y ofrecerle ofertas de productos/servicios antes de comenzar su traslado. Se podría también deducir, por ejemplo, que el usuario visita el centro de la ciudad los sábados por la tarde y ofrecerle anuncios de ofertas disponibles en dicho intervalo de tiempo. Estos son algunos de los antecedentes relacionados al proyecto presentado que fundamentan la factibilidad del desarrollo de un sistema de recomendación inteligente de anuncios. También se podría deducir el interés en cierto tipo de productos en base al tiempo de permanencia en determinadas tiendas.

3 Objetivos del proyecto

El objetivo general del proyecto consistía en diseñar un prototipo de asistente inteligente que permita vincular las necesidades de un usuario (capturadas de forma implícita) con las ofertas del portal de comercio electrónico Región20, ordenando las mismas por proximidad geográfica al usuario.

Los objetivos específicos del mismo consistieron en:

- Definir indicadores a extraer a partir de la información generada por el usuario en la navegación/búsqueda en el portal, e implementar algoritmo de extracción de dicha información.
- Definir e implementar un algoritmo de extracción de indicadores a partir de las publicaciones de productos / servicios y los usuarios que las realizan, poniendo especial énfasis en aspectos geoespaciales.
- Definir e implementar un algoritmo de perfilado de usuarios que permita, de forma inteligente, caracterizar sus necesidades en el mediano y corto plazo, a partir de los indicadores extraídos.
- Definir e implementar métricas de similitud de usuarios que permitan minimizar los problemas del “arranque en frío” del algoritmo de predicción.
- Definir e implementar un prototipo asistente inteligente que permita recomendar productos a los usuarios teniendo en cuenta sus necesidades y restricciones geoespaciales.
- Definir un framework de evaluación que permita medir el éxito en las recomendaciones.

4 Etapas del proyecto

El proyecto fue dividido en tres etapas.

Etapa 1: Extracción del conocimiento

En esta etapa se abordan los objetivos específicos 1 y 2: para este fin es necesario conocer la información contextual disponible.

- Documentación de la información disponible en el portal Region20 y posibles formas de acceso a la misma Definición de Metodología de Trabajo Análisis de factibilidad de los algoritmos existentes en la literatura
- Identificación de técnicas de recomendación utilizadas en el dominio
- Selección de indicadores aptos y algoritmo de extracción de perfiles de usuario y productos

Como resultado, se elaboró un informe técnico con un resumen detallado de la información disponible para la caracterización de los usuarios y de las publicaciones, un análisis de los algoritmos de perfilado de usuario existentes y un listado de aquellos que resulten más aptos para el problema.

Etapas 2: Perfilado inteligente del Usuario

En esta etapa se diseñaron e implementaron los perfiles de usuario, se utilizará un algoritmo de caracterización individual del usuario y métricas de similitud de usuarios que permitan realizar recomendaciones a usuarios sin suficiente información individual. Esta etapa abarca los objetivos específicos 3 y 4.

- Análisis de técnicas de perfilado individual de un usuario para la recomendación de productos. Se pondrá un énfasis especial en la información general y en la de corto plazo, es decir interés actual sobre una compra.
- Análisis de métricas de similitud entre usuarios. Sobre la base de los indicadores individuales realizar un análisis de técnicas de clustering que permitan una mejor caracterización del usuario nuevo del sistema.
- Diseño de algoritmo de perfilado inteligente de usuario. A partir de los análisis previos, diseñar un algoritmo que permita perfilar a un usuario a partir de la información disponible.
- Implementación del prototipo del perfilador inteligente de usuarios .

Como resultado, se elaboró un informe y se desarrolló el prototipo de algoritmo perfilador de usuarios.

Etapas 3: Recomendación de Productos

A partir de la información de las etapas anteriores, se diseñó e implementó un prototipo de algoritmo de recomendación de anuncios disponibles en el portal para los usuarios del mismo. Se tendrá en cuenta el perfil del usuario así como la información geoespacial adquirida previamente. Se abordan en esta etapa los objetivos específicos 5 y 6 .

- Diseño e implementación de un recomendador de publicaciones para los usuarios.
- Diseño de Framework para la validación de recomendaciones.
- Prueba del prototipo generado en laboratorio.
- Prueba del prototipo en condiciones reales .
- Experimentación con aplicaciones móviles geoposicionadas.

- Ajuste del prototipo acorde a los resultados experimentales .
- Discusión de los resultados obtenidos, conclusiones y colaboraciones futuras.

5 Resultados alcanzados

A partir del desarrollo del proyecto se fueron alcanzando diversos objetivos iniciales y los que surgieron de la colaboración entre las partes del proyecto.

Perfil del usuario

Con el análisis de la información disponible en el portal se estableció un perfil individual para cada usuario. Dicho perfil puede actualizarse a medida que se establecen nuevas interacciones. El perfil registra las interacciones directas e indirectas. A partir del clasificador desarrollado también se registran las potenciales categorías de publicaciones en las que estaba interesado el usuario en el momento de hacer una búsqueda. La jerarquización de las características permite desarrollar futuras aplicaciones que permitan llegar de forma personalizada al usuario. El perfil permite:

- Obtener el interés histórico de un usuario
- Obtener el interés de un usuario dada una ventana de tiempo
- Dada una publicación, determinar si la misma resulta de interés o no a un usuario

Clasificador de publicaciones

Se propuso un enfoque de un meta clasificador de 3 niveles que permite obtener las posibles categorías a partir de un conjunto de palabras. Este clasificador, el cual fue evaluado y probado exhaustivamente, permite clasificar las interacciones de un usuario cuando busca en el portal, así como también dado el título de una nueva publicación categorizarla de manera automática. Este último aspecto resultó de suma importancia para la empresa.

Aplicaciones dentro del dominio

A partir de las diversas reuniones surgieron diferentes potenciales aplicaciones del proyecto realizado. Actualmente se está finalizando con el proceso de migración del sitio a PHP 8. El próximo hito (planificado para el 2022) es la mejora en la publicación de avisos, donde estará la integración con el categorizador presentado y más adelante la obtención del perfil de usuarios.

- **Mailing personalizado.** El portal realiza campañas de mailing esporádicas con los últimos avisos o acorde a fechas importantes, tales como día del niño, día de la madre, etc. A partir de los objetivos alcanzados, es posible armar el mail de forma personalizada, es decir, dado un usuario incluir una serie potencial de artículos que le resulten de su interés primario.

- **Últimos artículos publicados.** Cuando un nuevo artículo es aprobado para su publicación es posible obtener un listado de potenciales usuarios que son de su interés.
- **Carrousel personalizado de artículos.** Actualmente, en la navegación del sitio se muestran artículos relacionados al artículo que el usuario está viendo y no artículos de su potencial interés. Con la nueva funcionalidad, es posible que dicho carrousel de artículos sea armado en base a los intereses de los usuarios.

Categorización de nuevos artículos. En este caso es una funcionalidad que se encuentra en desarrollo y utiliza como pilar el clasificador desarrollado. Esta funcionalidad permite dado un título de artículo generar automáticamente las categorías asociadas al mismo. La categorización de la predicción en tres etapas permite que la funcionalidad se integre de forma natural, ya que al usuario se le seleccionarán de forma automática las categorías predichas, pudiendo cambiarlas si lo desea. Este clasificador permitirá establecer un criterio común entre los moderadores, lo que hará que con el tiempo todas las publicaciones converjan de forma natural.

6 Conclusiones

Para la empresa, el desarrollo le permite incorporar el uso de técnicas de Inteligencia Artificial a su portal de ventas online Region20. Los resultados de este proyecto, le permitirán incorporar a futuro nuevas funcionalidades a su sitio. Se podrá, entre otras cuestiones: a) realizar marketing de manera personalizada y utilizando mejor los recursos disponibles. El recomendador de publicaciones y el perfilado inteligente de usuarios es una herramienta muy potente a la hora de llegar al usuario de forma personalizada y eficaz. b) mejorar la navegación del sitio, en un futuro es posible que la interfaz se adapte al perfil del usuario mejorando plenamente su experiencia en el sitio. c) generación de aplicaciones móviles efectivas y que faciliten la adopción por parte de los usuarios.

Para el grupo de trabajo consistió en un puntapié fundamental para la transferencia de conocimiento de forma efectiva a la empresa. El grupo se beneficia con la posibilidad práctica de materializar algoritmos investigados, acceso a información real que permita validar nuevas técnicas de personalización. Asimismo, el presente proyecto es un inicio en la relación entre la empresa y el grupo de trabajo que podrá consolidarse a partir de nuevos proyectos que permitan transferir conocimiento de forma efectiva.

Una valiosa lección aprendida de las interacciones y reuniones constantes entre las partes es como el proyecto cobra vida y toma direcciones inesperadas. De las reuniones surgió la necesidad de mejorar el tiempo de publicación para incentivar que los usuarios publiquen sus artículos. En este aspecto se vio cómo elegir la categoría era una traba importante del proceso. Cuando se desarrollaba el perfil del usuario se vio la potencialidad de este desarrollo (categorizar las publicaciones acorde a su título

y descripción) en mejorar el proceso de publicación. Esto llevó a dedicarle mucho más tiempo ya que esto permitiría simplificar el proceso para el usuario e incluso para los revisores que debían aprobar o no publicaciones.

Referencias

- [1] C. Mihai, C. Popa and P. M. Mircea, "Load Profiling For Gas Stations Using Cluster Techniques", IEEE International Power Electronics and Motion Control Conference (PEMC), Varna, pp. 1041-1048, 2016.
- [2] J. Yang, C. Liu, M. Teng, M. Liao and H. Xiong, "Buyer Targeting Optimization: A Unified Customer Segmentation Perspective", IEEE International Conference on Big Data (Big Data), Washington, DC, pp. 1262- 1271, 2016.
- [3] A. D. Rachid, A. Abdellah, B. Belaid, L. Rachid, "Clustering Prediction Techniques in Defining and Predicting Customers Defection: The Case of E-Commerce Context", International Journal of Electrical and Computer Engineering (IJECE), ISSN: 2088-8708. Vol 8, No 4: August, 2018.
- [4] Girish S, Ramamurthy B, Senthilnathan T, "Mining the Web Data for Classifying and Predicting Users' Requests", International Journal of Electrical and Computer Engineering (IJECE), ISSN: 2088-8708. Vol 8, No 4, 2018.
- [5] Yuan, S.-T., & Tsao, Y. W. (2003). A recommendation mechanism for contextualized mobile advertising. *Expert Systems with Applications*, 24(4), 399–414.
- [6] Zaim, H., Haddi, A., & Ramdani, M. (2019). A novel approach to dynamic profiling of e-customers considering click stream data and online reviews. *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 9, p. 602.
- [7] Molitor, D., Reichhart, P., Spann, M., & Ghose, A. (2019). Measuring the Effectiveness of Location-Based Pull Advertising: A Randomized Field Experiment. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2645281>
- [8] Vallejos, S., and B. Caimmi. 2016. "Geolocalización E Incidentes de Tránsito a Partir Del Análisis de Sentencias Expresadas En Lenguaje Natural Extraídas de Redes Sociales." Edited by A. Soria L. Berdún. Ingeniería de Sistemas, Facultad de Ciencias Exactas, UNCPBA.
- [9] Caimmi, B, Vallejos S., Berdun L. Soria A., Amandi A, and Campo M. 2016. "Detección de Incidentes de Tránsito En Twitter." In 2016 IEEE Biennial Congress of Argentina (ARGENCON). <https://doi.org/10.1109/argencon.2016.7585327>.
- [11] Vallejos S., Caimmi, B, Alonso D., Vallejos S., Berdun L. and Soria A. 2018. "Comparing detection and disclosure of traffic incidents in social networks: an intelligent approach based on Twitter vs. Waze" *Inteligencia Artificial, Iberoamerican Journal of Artificial Intelligence (Iberamia)* Vol 21 No 61 ISSN 1988-3064
- [11] D’Cristóforo, M. S., and A. H. Giannoni. 2016. "Asistencia Inteligente En La Planificación Personalizada de La Movilidad Urbana." In *Proceedings de EST'16, Concurso de Trabajos Estudiantiles*. Directores: A. Soria y L. Berdun.
- [12] Vallejos, S.; Armentano, M.; Berdun, L.. 2019. "TourWithMe: Recommending peers to visit attractions together" En *Proceedings of RecTour, Workshop on Recommender Systems in Tourism. ACM International Conference on Recommender Systems*
- [13] Nigro, O y González Císaro, S. 2018. Plan director de transporte urbano y la movilidad en la ciudad de Tandil. Proyecto de cooperación Municipio de Tandil - Universidad Nacional del Centro de la Provincia de Buenos Aires.